A simple method to balanced Procrustes problem with one special constraint

Zhongyang Yuan
Zhejiang Gongshang University
College of Statistics and Mathematics
Hangzhou, 310018
P. R. China
yuanzhongyang@126.com

Yuyang Qiu
Zhejiang Gongshang University
College of Statistics and Mathematics
Hangzhou, 310018
P. R. China
YuyangQiu@gmail.com

Abstract: The balanced Procrustes problem with $X^T = sX$ and $XX^T = aX + bI_n$ constraints are considered. By one time eigenvalue decomposition or real Schur decomposition of the matrix product generated by the matrices A and B, the constrained solutions are constructed simply. Similar strategy is applied to the problem with the corresponding P-commuting constraints with given symmetric matrix P. And the methods are also suitable for the least squares problem of the extended equations $AX = B,\ XC = D$ with the same constraints. Numerical examples are presented to show the efficiency of the proposed methods.

Key–Words: Balanced Procrustes problem, Eigenvalue decomposition, Real Schur decomposition, Matrix product, Constrained solutions, Numerical examples

1 Introduction

Given matrices $A,\ B \in \mathbb{R}^{m\times n}$, we consider the following matrix equation

$$AX = B \tag{1}$$

and its corresponding least squares problem (balanced Procrustes problem)

$$\min_X \|AX - B\|_F \tag{2}$$

with $X^T = sX$ and

$$XX^T = aX + bI_n \tag{3}$$

constraints, where the unknown matrix $X \in \mathbb{R}^{n\times n}$ and scalars $s,\ a,\ b$ satisfy

$$\begin{array}{ll} a^2 + 4b > 0, \ \text{ for } \ s = 1; \\ a = 0, \ \ b > 0, \ \text{ for } \ s = -1. \end{array} \tag{4}$$

And, n is even if $s = -1$.

In the passed decades, the equation (1) with different constraints, together with its corresponding least squares problem (2), has been interested widely, more and more people have been attracted in this topic [1, 2, 3, 4, 6, 7, 8, 9, 10, 14, 16, 17, 18, 19]. For the equation (1) with reflexive and anti-reflexive constraints, Peng and Hu transformed the constrained equation to two independent unconstrained ones, with which they obtained the existence condition and detailed structure of constrained solutions [13]. For the equation (1) with $X = X^T$ and $XX^T = I_n$ constraints, Meng and Hu constructed its general solution by four times singular value decompositions and one eigenvalue decomposition of related matrices [12], Qiu and Wang proposed a more efficient algorithm to obtained the solution by one eigenvalue decomposition in [15]. For the equation (1) with $X = -X^T$ and $XX^T = I_n$ constraints, Meng, Hu and Zhang obtained its general solution in terms of C-S decomposition and Schur decomposition [11].

In this paper, we consider the balanced Procrustes problem (2) with constraints (3), our ideas are based on the following observations:

1) If the equation (1) is consistent, then its solutions must be included in that of the problem (2) with zero residual. Hence, the constrained equation (1) can be regarded as the special case of the corresponding balanced Procrustes problem (2).

2) By selecting the suitable numbers of $a,\ b$ in (3), we can get the constrained problems in [12] and [11] respectively. From this point, our problem maybe is more general.

3) The selection of parameters $a,\ b$ in (3) is to guarantee the unknown matrix X has two clusters of different eigenvalues.

Note that

$$\begin{aligned}\|AX-B\|_F^2 &= \operatorname{trace}(AX-B)^T(AX-B) \quad (5)\\ &= \operatorname{trace}(A^TAXX^T) - 2\operatorname{trace}(X^TA^TB) + \|B\|_F^2.\end{aligned}$$

In [5], Golub & Van Loan constructed the general solution of the problem (2) with orthogonal constraint by a simple method: with the constraint, the equivalent form of (5) is transformed to minimize a trace composed by a linear mapping of unknown X, and the constrained solution is obtained by one time SVD of the matrix product of A and B. Motivated by this idea, we consider the problem (2) with constraints (3) with the same strategy. That is, the constrained problem can be simplified as the following minimizing problem :

$$\min_{X^T=sX,\ XX^T=aX+bI_n} \operatorname{trace}(WX) \quad (6)$$

with $W = aA^TA - 2sA^TB$. Therefore, we ask for the constrained solutions only by one time eigenvalue decomposition or the real Schur decomposition of the matrix product generated by A and B. Compared with the existing methods in [11, 12], we have the following improvements:

1) For the equation (1) with symmetric orthogonal constraint which was considered in [12], we continue along the same line of research of [15] by Qiu and Wang, and generalize the constraint from symmetric orthogonality to (3).

2) For the equation (1) with skew-symmetric orthogonal constraint which was considered in [11], we regard it as the special case of the problem (2) with the same constraint and obtain its constrained solution only by one real-Shur decompositions. By computing again and again, we find our methods maybe are more efficient (One can turn to Example 1 of section 5 for details).

3) Our conclusions can be generalized to the corresponding P-commuting constrained least squares problems and the extended matrix equations $AX = B,\ XC = D$ with the same constraints.

The rest of paper is organized as follows. We consider the problem (2) with constraints (3) in Section 2. The corresponding P-commuting constrained problems are discussed in Section 3. In Section 4, as a special case of the problem (2), we consider the least squares problem of the equations $AX = B,\ XC = D$ with the same constraints. Numerical examples are given in Section 5 to display the efficiency of the algorithms.

Notations. In this paper, $\mathbb{R}^{m\times n}$ denotes the space of real $m \times n$ matrix. For any matrix $X = (x_{ij}) \in \mathbb{R}^{m\times n}$, $\operatorname{trace}(X)$ is its trace, $\|\cdot\|_F$ is the Frobenius norm of matrix. If X is a square matrix with order n,

$$D_X = \operatorname{diag}(x_{11}, \ldots, x_{nn}) \quad (7)$$

is the diagonal matrix composed by the diagonal elements of X. We also denote by I_n the identity matrix with order n. The matrix $O_{m\times n}$ is $m \times n$ zero matrix, O_n is $n \times n$ zero matrix and sign(.) refers to the symbolic function, respectively. Moreover, for $b > 0$, we denote

$$\tilde{D}_b = \begin{pmatrix} 0 & \sqrt{b} \\ -\sqrt{b} & 0 \end{pmatrix} \quad (8)$$

2 Solutions to the balanced Procrustes problem (2) with (3) constraints

With constraints (3),

$$\begin{aligned}&\operatorname{trace}(WX) = \operatorname{trace}(X^TW^T)\\ =\ &\operatorname{trace}(sXW^T) = \operatorname{trace}(sW^TX),\end{aligned}$$

then

$$\operatorname{trace}(WX) = \frac{1}{2}\operatorname{trace}((W + sW^T)X).$$

Hence, the problem (2) with constraints (3) is equivalent to

$$\min_{X^T=sX,\ XX^T=aX+bI_n} \operatorname{trace}((W + sW^T)X). \quad (9)$$

2.1 Case $s = 1$

If $s = 1$, the matrix $(W + sW^T)$ is symmetric. The condition (3) is equivalent to

$$X^2 = aX + bI_n. \quad (10)$$

Let λ be the eigenvalue of X, it satisfies

$$\lambda^2 - a\lambda - b = 0. \quad (11)$$

Denote by $\Delta = a^2 + 4b$. With (4), we have $\Delta > 0$, which implies equation (11) has two different roots

$$\lambda_{1,2} = \frac{a \pm \sqrt{\Delta}}{2}. \quad (12)$$

So the constrained solution X has two clusters of eigenvalues, we denote them by

$$\lambda_1 > \lambda_2.$$

Let an eigenvalue decomposition of the symmetric matrix $W + W^T$ be

$$W + W^T = Q \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & O_{n-r_1-r_2} \end{pmatrix} Q^T, \quad (13)$$

where

$$\begin{array}{l} \Lambda_1 = \mathrm{diag}(\lambda_1^+, \ldots, \lambda_{r_1}^+), \\ \Lambda_2 = \mathrm{diag}(\lambda_1^-, \ldots, \lambda_{r_2}^-), \\ r_1 + r_2 = \mathrm{rank}(W + W^T), \\ \lambda_i^+ > 0,\ \lambda_j^- < 0, \\ i = 1, \ldots, r_1,\ j = 1, \ldots, r_2, \\ Q^T Q = I_n. \end{array}$$

Set $\tilde{X} = Q^T X Q$ and denote its element by

$$\tilde{X} = (\tilde{x}_{ij}),\ i, j = 1, \ldots, n. \quad (14)$$

The constraint $XX^T = aX + bI_n$ implies

$$\tilde{X}\tilde{X}^T = a\tilde{X} + bI_n,$$

so

$$\sum_{k=1}^{n} \tilde{x}_{ki}^2 = a\tilde{x}_{ii} + b,\ i = 1, 2, \ldots, n. \quad (15)$$

Therefore

$$\tilde{x}_{ii}^2 \leq a\tilde{x}_{ii} + b,$$

that is

$$\lambda_2 \leq \tilde{x}_{ii} \leq \lambda_1,\ \ i = 1, \ldots, n. \quad (16)$$

Partition $\tilde{X}$ by

$$\tilde{X} = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}, \quad (17)$$

then

$$\mathrm{trace}((W + W^T)X) = \mathrm{trace}\left(\Lambda_1 X_{11} + \Lambda_2 X_{22}\right).$$

The question (6) is equivalent to

$$\min_{X_{11}, X_{22}} \mathrm{trace}\left(\Lambda_1 X_{11} + \Lambda_2 X_{22}\right). \quad (18)$$

With inequality (16), it is not difficult to verify

$$D_{X_{11}} = \lambda_2 I_{r_1},\ D_{X_{22}} = \lambda_1 I_{r_2}.$$

Together with (15), we have

$$\begin{array}{l} X_{ij} = 0,\ i, j = 1, 2, 3,\ \ i \neq j; \\ X_{11} = \lambda_2 I_{r_1},\ X_{22} = \lambda_1 I_{r_2}. \end{array}$$

Thus, we have the following theorem.

Theorem 1 *Denote by $W = aA^TA - 2A^TB$ and suppose an eigenvalue decomposition of the matrix $W + W^T$ is*

$$W + W^T = Q \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & O_{n-r_1-r_2} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} \Lambda_1 = \mathrm{diag}(\lambda_1^+, \ldots, \lambda_{r_1}^+), \\ \Lambda_2 = \mathrm{diag}(\lambda_1^-, \ldots, \lambda_{r_2}^-), \\ r_1 + r_2 = \mathrm{rank}(W + W^T), \\ \lambda_i^+ > 0,\ \lambda_j^- < 0, \\ i = 1, \ldots, r_1,\ j = 1, \ldots, r_2, \\ Q^T Q = I_n. \end{array}$$

The general solutions of problem (2) with

$$X^T = X \ \text{and}\ XX^T = aX + bI_n$$

constraints are

$$X = Q \begin{pmatrix} \lambda_2 I_{r_1} & & \\ & \lambda_1 I_{r_2} & \\ & & G \end{pmatrix} Q^T, \quad (19)$$

where $\lambda_1,\ \lambda_2$ are determined by (12), the matrix $G \in \mathbb{R}^{n-r_1-r_2 \times n-r_1-r_2}$ satisfies

$$G^T = G,\ GG^T = aG + bI_{n-r_1-r_2}$$

and $a^2 + 4b > 0$.

Now we consider the two special cases of Theorem 1. For $a = 0,\ b = 1$, the constraints are

$$X^T = X,\ XX^T = I_n, \quad (20)$$

that is the symmetric orthogonal constraint, we have the following corollary.

Corollary 2 *Denote by $W = -2A^TB$ and suppose an eigenvalue decomposition of the matrix $W + W^T$ is*

$$W + W^T = Q \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & O_{n-r_1-r_2} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} \Lambda_1 = \mathrm{diag}(\lambda_1^+, \ldots, \lambda_{r_1}^+), \\ \Lambda_2 = \mathrm{diag}(\lambda_1^-, \ldots, \lambda_{r_2}^-), \\ r_1 + r_2 = \mathrm{rank}(W + W^T), \\ \lambda_i^+ > 0,\ \lambda_j^- < 0, \\ i = 1, \ldots, r_1,\ j = 1, \ldots, r_2, \\ Q^T Q = I_n. \end{array}$$

The general solutions of problem (2) with symmetric orthogonal constraint (see (20)) are

$$X = Q \begin{pmatrix} -I_{r_1} & & \\ & I_{r_2} & \\ & & G \end{pmatrix} Q^T, \tag{21}$$

where $G \in \mathbb{R}^{n-r_1-r_2 \times n-r_1-r_2}$ *satisfies*

$$G^T = G,\ GG^T = I_{n-r_1-r_2}.$$

For $a = 1,\ b = 0$, the constraints are

$$X^T = X,\ XX^T = X, \tag{22}$$

that is the symmetric idempotent constraint. And we also have the following corollary.

Corollary 3 *Denote by* $W = A^T A - 2A^T B$ *and suppose an eigenvalue decomposition of the matrix* $W + W^T$ *is*

$$W + W^T = Q \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & O_{n-r_1-r_2} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} \Lambda_1 = \mathrm{diag}(\lambda_1^+, \ldots, \lambda_{r_1}^+), \\ \Lambda_2 = \mathrm{diag}(\lambda_1^-, \ldots, \lambda_{r_2}^-), \\ r_1 + r_2 = \mathrm{rank}(W + W^T), \\ \lambda_i^+ > 0,\ \lambda_j^- < 0, \\ i = 1, \ldots, r_1,\ j = 1, \ldots, r_2, \\ Q^T Q = I_n. \end{array}$$

The general solutions of problem (2) with symmetric idempotent constraint (see (22)) are

$$X = Q \begin{pmatrix} O_{r_1} & & \\ & I_{r_2} & \\ & & G \end{pmatrix} Q^T, \tag{23}$$

where $G \in \mathbb{R}^{n-r_1-r_2 \times n-r_1-r_2}$ *satisfies*

$$G^T = G,\ GG^T = G.$$

2.2 Case $s = -1$

For $s = -1$, $W = A^T B$ and $W - W^T$ is skew-symmetric, we consider

$$\min_{X^T = -X,\ XX^T = bI_n} \mathrm{trace}((W - W^T)X). \tag{24}$$

Let a real Schur decomposition of the matrix $W - W^T$ be

$$W - W^T = Q \begin{pmatrix} C & \\ & O_{n-2r} \end{pmatrix} Q^T, \tag{25}$$

where

$$\begin{array}{l} C = \mathrm{diag}(C_1, \ldots, C_r), \\ C_i = \begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix},\ c_i > 0, \\ 2r = \mathrm{rank}(W - W^T),\ i = 1, \ldots, r. \\ Q^T Q = I_n. \end{array}$$

Set $\tilde{X} = Q^T X Q$ and denote

$$\tilde{X} = (\tilde{x}_{ij}), \quad i, j = 1, 2, \ldots, n.$$

We have

$$((W - W^T)X)_{2i-1,2i-1} = c_i \tilde{x}_{2i,2i-1}$$

and

$$((W - W^T)X)_{2i,2i} = -c_i \tilde{x}_{2i-1,2i}$$

with $i = 1, 2, \ldots, r$. The constraints $X^T = -X$ and $XX^T = bI_n$ imply

$$\tilde{X}^T = -\tilde{X},\ \tilde{X}\tilde{X}^T = bI_n,$$

so

$$\sum_{k=1}^{n} \tilde{x}_{ki}^2 = b,\ i = 1, 2, \ldots, n. \tag{26}$$

To solve the equation (24), we have

$$\tilde{x}_{2i-1,2i} = \sqrt{b},\ i = 1, 2, \ldots, r. \tag{27}$$

Together with (26), the following equalities hold:

$$\tilde{X} = \begin{pmatrix} \tilde{D}_b & & & \\ & \ddots & & \\ & & \tilde{D}_b & \\ & & & G \end{pmatrix},$$

where $\tilde{D}_b$ is defined by (8), and G is arbitrary. Thus, the following theorem holds.

Theorem 4 *Denote by* $W = A^T B$ *and suppose a real Schur decomposition of* $W - W^T$ *is*

$$W - W^T = Q \begin{pmatrix} C & \\ & O_{n-2r} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} C = \mathrm{diag}(C_1, \ldots, C_r), \\ C_i = \begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix},\ c_i > 0, \\ 2r = \mathrm{rank}(W - W^T),\ i = 1, \ldots, r, \\ Q^T Q = I_n. \end{array}$$

The general solutions of problem (2) with

$$X = -X^T \;\; and \;\; XX^T = bI_n$$

constraints are

$$X = Q \begin{pmatrix} \tilde{D}_b & & & \\ & \ddots & & \\ & & \tilde{D}_b & \\ & & & G \end{pmatrix} Q^T,$$

where $\tilde{D}_b$ is defined by (8) and $G \in \mathbb{R}^{n-2r\times n-2r}$ satisfies

$$G^T = -G, \; G^T G = bI_n.$$

Remark 5 *1) In this problem, $a = 0$ is based on the following observation:*

$$\begin{aligned} XX^T &= (XX^T)^T \\ &= (aX + bI_n)^T \\ &= -aX + bI_n, \end{aligned}$$

which implies that $a = 0$.

2) For $b = 1$, the constraint is skew-symmetric orthogonal.

3 Solutions to the balanced Procrustes problem with corresponding P-commuting constraints

In this section, we generalize the constraints to their corresponding P-commuting constraints with given symmetric matrix P, that is, we want to consider the problem (2) with

$$PX = XP, \; X^T = sX \;\; \text{and} \;\; XX^T = aX + bI_n$$

constraints.

Let an eigenvalue decomposition of P be

$$P = V\mathrm{diag}(\bar{\lambda}_1 I_{k_1}, \cdots, \bar{\lambda}_p I_{k_p})V^T, \tag{28}$$

where $V^T V = I_n$, k_i is the multiples of eigenvalues $\bar{\lambda}_i$ satisfying $\Sigma_{i=1}^p k_i = n$. Note that a matrix X satisfies

$$PX = XP$$

if and only if

$$X = V\mathrm{diag}(X_1, \cdots, X_p)V^T, \tag{29}$$

where

$$X_i \in \mathbb{R}^{k_i\times k_i}, \; i = 1, \ldots, p.$$

If X is further required to satisfy $X^T = sX$ and $XX^T = aX + bI_n$, then all $\{X_i\}_{i=1}^p$ in (29) should satisfy the same constraints too, that is,

$$\begin{aligned} X_i^T &= sX_i, \; X_i X_i^T = aX_i + bI_{k_i}, \\ i &= 1, \ldots p. \end{aligned}$$

Denote by $\tilde{W} = V^T W V$ with $W = aA^T A - 2sA^T B$. And partition the matrix

$$\tilde{W} = (W_{ij})$$

conforming to (29). So the constrained solutions are represented by $X = V\mathrm{diag}(X_1, \cdots, X_p)V^T$, where X_i satisfies

$$\min_{X_i^T = sX_i, \; X_i X_i^T = aX_i + bI_{k_i}} \mathrm{trace}(X_i(W_{ii} + sW_{ii}^T)),$$
$$i = 1, \ldots, p,$$

which can be solved by Theorem 1 and Theorem 4. Hence, we have the following theorems.

Theorem 6 *Denote by $\tilde{W} = V^T W V$ with $W = aA^T A - 2A^T B$, and the matrix V is determined by (28). We partition the matrix $\tilde{W} = (W_{ij})$ conforming to (29), where*

$$W_{ij} \in \mathbb{R}^{k_i\times k_j}.$$

Let an eigenvalue decomposition of the matrix $W_{ii} + W_{ii}^T$ be

$$W_{ii} + W_{ii}^T = Q_i \begin{pmatrix} \Lambda_1^{(i)} & & \\ & \Lambda_2^{(i)} & \\ & & O_{n-r_1^{(i)}-r_2^{(i)}} \end{pmatrix} Q_i^T,$$

where

$$\begin{aligned} &\Lambda_1^{(i)} = \mathrm{diag}(\lambda_1^{(i)+}, \ldots, \lambda_{r_1^{(i)}}^{(i)\,+}), \\ &\Lambda_2^{(i)} = \mathrm{diag}(\lambda_1^{(i)-}, \ldots, \lambda_{r_2^{(i)}}^{(i)\,-}), \\ &r_1^{(i)} + r_2^{(i)} = \mathrm{rank}(W_{ii} + W_{ii}^T), \\ &\lambda_j^{(i)+} > 0, \; \lambda_k^{(i)-} < 0, \\ &j = 1, \ldots, r_1^{(i)}, \; k = 1, \ldots, r_2^{(i)}. \\ &Q_i^T Q_i = I_{k_i}. \end{aligned}$$

The solutions to the problem (2) with

$$PX = XP, \; X^T = X \;\; and \;\; XX^T = aX + bI_n$$

constraints are

$$X = V\mathrm{diag}(X_1, \cdots, X_p)V^T,$$

where

$$X_i = Q_i \begin{pmatrix} \lambda_2 I_{r_1^{(i)}} & & \\ & \lambda_1 I_{r_2^{(i)}} & \\ & & G_i \end{pmatrix} Q_i^T,$$

$\lambda_1,\ \lambda_2$ *are determined by (12), and the matrix* $G_i \in \mathbb{R}^{(k_i - r_1^{(i)} - r_2^{(i)}) \times (k_i - r_1^{(i)} - r_2^{(i)})}$ *satisfies*

$$G_i^T = G_i,\ G_i G_i^T = aG_i + bI_{k_i - r_1^{(i)} - r_2^{(i)}}.$$

Theorem 7 *Denote by* $\tilde{W} = V^T W V$ *with* $W = A^T B$, *and the matrix* V *is determined by (28). We partition the matrix* $\tilde{W} = (W_{ij})$ *conforming to (29), where* $W_{ij} \in \mathbb{R}^{k_i \times k_j}$. *Let a real Schur decomposition of the matrix* $W_{ii} - W_{ii}^T$ *be*

$$W_{ii} - W_{ii}^T = Q_i \begin{pmatrix} C^{(i)} & \\ & O_{k_i - r_i} \end{pmatrix} Q_i^T, \qquad (30)$$

where

$$\begin{array}{l} C^{(i)} = \mathrm{diag}(C_1^{(i)}, \ldots, C_{r_i}^{(i)}), \\ C_j^{(i)} = \begin{pmatrix} 0 & c_j^{(i)} \\ -c_j^{(i)} & 0 \end{pmatrix}, \\ 2r_i = \mathrm{rank}(W_{ii}^T - W_{ii}), \\ c_j^{(i)} > 0,\ j = 1, \ldots, r_i, \\ Q_i^T Q_i = I_{k_i}. \end{array}$$

The solutions to the problem (2) with

$$PX = XP,\ X^T = -X \ \ \text{and} \ \ XX^T = bI_n$$

constraints are

$$X = V \mathrm{diag}(X_1, \cdots, X_p) V^T$$

with

$$X_i = Q_i \mathrm{diag}(\tilde{D}_b, \ldots, \tilde{D}_b, G_i) Q_i^T,$$

where $\tilde{D}_b$ *is defined by (8) and* $G_i \in \mathbb{R}^{k_i - 2r_i \times k_i - 2r_i}$ *satisfies*

$$G_i^T = -G_i,\ G_i G_i^T = bI_{k_i - 2r_i}.$$

4 The least squares problem of one extended matrix equations with the same constraints

In this section, we consider the least squares problems of the extended matrix equations

$$AX = B,\ XC = D \qquad (31)$$

with the same constraints, where matrices

$$A,\ B \in \mathbb{R}^{m \times n},\ \ C,\ D \in \mathbb{R}^{n \times p}.$$

That is, we want to consider the following least squares problem

$$\min_X \| \begin{pmatrix} AX - B \\ C^T X^T - D^T \end{pmatrix} \|_F \qquad (32)$$

with

$$X^T = sX \ \ \text{and} \ \ X^T X = aX + bI_n$$

constraints. Note that (32) is equivalent to

$$\min_X \|\tilde{A}X - \tilde{B}\|_F \qquad (33)$$

with

$$\tilde{A} = \begin{pmatrix} A \\ sC^T \end{pmatrix}, \tilde{B} = \begin{pmatrix} B \\ D^T \end{pmatrix}.$$

Hence, their constrained least squares solutions can been obtained in terms of Theorem 1, Theorem 4, Theorem 6 and Theorem 7 only by replacing $A = \tilde{A}$ and $B = \tilde{B}$ in W, respectively. Therefore, we have the following theorems.

Theorem 8 *Denote by*

$$W = a(A^T A + CC^T) - 2(A^T B + CD^T),$$

and suppose an eigenvalue decomposition of the matrix $W + W^T$ *is*

$$W + W^T = Q \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & O_{n - r_1 - r_2} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} \Lambda_1 = \mathrm{diag}(\lambda_1^+, \ldots, \lambda_{r_1}^+), \\ \Lambda_2 = \mathrm{diag}(\lambda_1^-, \ldots, \lambda_{r_2}^-), \\ r_1 + r_2 = \mathrm{rank}(W + W^T), \\ \lambda_i^+ > 0,\ \lambda_j^- < 0, \\ i = 1, \ldots, r_1,\ j = 1, \ldots, r_2, \\ Q^T Q = I_n. \end{array}$$

The general solutions of problem (2) with

$$X^T = X \ \ \text{and} \ \ XX^T = aX + bI_n$$

constraints are

$$X = Q \begin{pmatrix} \lambda_2 I_{r_1} & & \\ & \lambda_1 I_{r_2} & \\ & & G \end{pmatrix} Q^T, \qquad (34)$$

where $\lambda_1,\ \lambda_2$ *are determined by (12), and the matrix* $G \in \mathbb{R}^{n - r_1 - r_2 \times n - r_1 - r_2}$ *satisfies*

$$G^T = G,\ GG^T = aG + bI_{n - r_1 - r_2}.$$

Theorem 9 *Denote by $\tilde{W} = V^TWV$ with*

$$W = a(A^TA + CC^T) - 2(A^TB + CD^T),$$

and the matrix V is determined by (28). We partition the matrix $\tilde{W} = (W_{ij})$ conforming to (29), where $W_{ij} \in \mathbb{R}^{k_i\times k_j}$. Let an eigenvalue decomposition of the matrix $W_{ii} + W_{ii}^T$ be

$$W_{ii} + W_{ii}^T = Q_i \begin{pmatrix} \Lambda_1^{(i)} & & \\ & \Lambda_2^{(i)} & \\ & & O_{n-r_1^{(i)}-r_2^{(i)}} \end{pmatrix} Q_i^T,$$

where

$$\begin{array}{l} \Lambda_1^{(i)} = \mathrm{diag}(\lambda_1^{(i)+}, \ldots, \lambda_{r_1^{(i)}}^{(i)\;+}), \\ \Lambda_2^{(i)} = \mathrm{diag}(\lambda_1^{(i)-}, \ldots, \lambda_{r_2^{(i)}}^{(i)\;-}), \\ r_1^{(i)} + r_2^{(i)} = \mathrm{rank}(W_{ii} + W_{ii}^T), \\ \lambda_j^{(i)+} > 0,\ \lambda_k^{(i)-} < 0, \\ j = 1, \ldots, r_1^{(i)}, k = 1, \ldots, r_2^{(i)}. \\ Q_i^TQ_i = I_{k_i}. \end{array}$$

The solutions to the problem (2) with

$$PX = XP,\ X^T = X,\ \textit{and}\ XX^T = aX + bI_n,$$

constraints are

$$X = V\mathrm{diag}(X_1, \cdots, X_p)V^T,$$

where

$$X_i = Q_i \begin{pmatrix} \lambda_2 I & & \\ & \lambda_1 I & \\ & & G_i \end{pmatrix} Q_i^T,$$

$\lambda_1,\ \lambda_2$ are determined by (12), and the matrix $G_i \in \mathbb{R}^{(k_i-r_1^{(i)}-r_2^{(i)})\times(k_i-r_1^{(i)}-r_2^{(i)})}$ satisfies

$$G_i^T = G_i,\ G_iG_i^T = aG_i + bI_{k_i-r_1^{(i)}-r_2^{(i)}}.$$

Theorem 10 *Denote by $W = A^TB + CD^T$ and suppose a real Schur decomposition of the matrix $W - W^T$ be*

$$W - W^T = Q \begin{pmatrix} C & \\ & O_{n-2r} \end{pmatrix} Q^T,$$

where

$$\begin{array}{l} C = \mathrm{diag}(C_1, \ldots, C_r), \\ C_i = \begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix}, c_i > 0, \\ 2r = \mathrm{rank}(W - W^T),\ i = 1, \ldots, r, \\ Q^TQ = I_n. \end{array}$$

Then the general solutions of problem (2) with

$$X^T = -X\ \textit{and}\ XX^T = bI_n$$

constraints are

$$X = Q\mathrm{diag}(\tilde{D}_b, \ldots, \tilde{D}_b, G)Q^T, \tag{35}$$

where $\tilde{D}_b$ is defined by (8) and $G \in \mathbb{R}^{n-2r\times n-2r}$ satisfies

$$G^T = -G,\ GG^T = bI_{n-2r}.$$

Theorem 11 *Denote by $\tilde{W} = V^TWV$ with $W = A^TB + CD^T$, and the matrix V is determined by (28). We partition the matrix $\tilde{W} = (W_{ij})$ conforming to (29), where $W_{ij} \in \mathbb{R}^{k_i\times k_j}$. Let a Schur decomposition of the matrix $W_{ii} - W_{ii}^T$ be*

$$W_{ii} - W_{ii}^T = Q_i \begin{pmatrix} C^{(i)} & \\ & O_{k_i-r_i} \end{pmatrix} Q_i^T, \tag{36}$$

where

$$\begin{array}{l} C^{(i)} = \mathrm{diag}(C_1^{(i)}, \ldots, C_{r_i}^{(i)}), \\ C_j^{(i)} = \begin{pmatrix} 0 & c_j^{(i)} \\ -c_j^{(i)} & 0 \end{pmatrix}, \\ 2r_i = \mathrm{rank}(W_{ii} - W_{ii}^T), \\ c_j^{(i)} > 0,\ j = 1, \ldots, r_i, \\ Q_i^TQ_i = I_{k_i}. \end{array}$$

The solutions to the problem (2) with

$$PX = XP,\ X^T = -X\ \ \textit{and}\ \ X^TX = bI_n$$

constraints are

$$X = V\mathrm{diag}(X_1, \cdots, X_p)V^T$$

with

$$X_i = Q_i\mathrm{diag}(\tilde{D}_b, \ldots, \tilde{D}_b, G_i)Q_i^T,$$

where $\tilde{D}_b$ is defined by (8) and $G_i \in \mathbb{R}^{k_i-2r_i\times k_i-2r_i}$ satisfies

$$G_i^T = -G_i,\ G_iG_i^T = bI_{k_i-2r_i}.$$

5 Numerical Examples

In this section, we present some numerical examples to illustrate the effectiveness of our theorems. For simplicity, we set $m = n = p$ and restrict all matrices $A,\ B,\ C\ D,\ E \in \mathbb{R}^{n\times n}$.

All examples are performed by MATLAB 7.3 on a personal computer of the Intel Core2 Duo CPU T7250 with 2G memory.

Example 1. In this example, let $s = -1, a = 0, b = 1$ in (3). We compare the algorithm based on Theorem 4 and that in [11] for the equation (1) with

$$X^T = -X \text{ and } XX^T = I_n$$

constraints. The test matrices $A,\ B$ are constructed as follows:

$$\begin{aligned} A &= U_0 \mathrm{diag}(\sigma_1, \cdots, \sigma_r) V_0^T, \quad (37)\\ B &= U_0 \mathrm{diag}(\sigma_1, \cdots, \sigma_r)([V_0,\ \tilde{V}_0] G_0)^T, \end{aligned}$$

where

$$U_0 \in \mathbb{R}^{m\times r},\ V_0 \in \mathbb{R}^{n\times r},\ G_0 \in \mathbb{R}^{n\times r}$$

are column orthogonal, $\tilde{V}_0$ is the orthogonal complement of V_0, the principle submatrix with order r of G_0 is skew-symmetric, and the singular values $\sigma_i, i = 1 : r$ are uniformly distributed in the interval (0, 1).

It is not difficult to verify: if A and B are set by (37), then we have

$$BB^T = AA^T,\ AB^T = -BA^T,$$

so the equation (1) has a skew-symmetric orthogonal ($X^T = -X$ and $X^T X = I_n$ constraints) solution. Therefore, the residual error $\|AX - B\|_F$ of an optimal constrained least squares solution to the equation (1) should be zero.

In Table 1, let numbers $m = n$ be variant from 100 to 1000, $r = 50$, we list the experiment results for different matrix sizes. For different n, the residual error $\|AX - B\|_F$ in both algorithms can also reach 10^{-13}. Both the skew-symmetric error $\|X^T + X\|_F$ and orthogonal error $\|XX^T - I_n\|_F$ in Theorem 4 can reach 10^{-13}, however those in [11] is only 10^{-5}. And CPU time of Algorithm in [11] is almost 1.5 times as that of Theorem 4. From the computing results, our algorithm will be better in skew-symmetric error and orthogonal error.

Table 1: The algorithm based on Theorem 4 and that in [11] for the equation (1) with $X^T = -X$, and $XX^T = I_n$ constraint

n	algorithm	CPU(s)	$\|AX - B\|_F$
100	Theorem 4	0.031	1.03*10^{-13}
	algorithm in [11]	0.078	2.32* 10^{-13}
200	Theorem 4	0.20	1.91* 10^{-13}
	algorithm in [11]	0.27	4.94*10^{-13}
300	Theorem 4	0.82	2.05* 10^{-13}
	algorithm in [11]	1.11	8.83* 10^{-13}
500	Theorem 4	3.81	3.57* 10^{-13}
	algorithm in [11]	5.34	2.14* 10^{-13}
700	Theorem 4	11.36	4.75*10^{-13}
	algorithm in [11]	16.26	2.31* 10^{-13}
1000	Theorem 4	31.33	6.27* 10^{-13}
	algorithm in [11]	45.32	4.77* 10^{-13}
n	algorithm	$\|X^T + X\|_F$	$\|XX^T - I_n\|_F$
100	Theorem 4	2.69* 10^{-15}	3.09* 10^{-13}
	algorithm in [11]	8.85* 10^{-06}	1.11* 10^{-06}
200	Theorem 4	3.96* 10^{-15}	2.69* 10^{-15}
	algorithm in [11]	1.33* 10^{-06}	1.69* 10^{-06}
300	Theorem 4	5.11* 10^{-15}	5.58* 10^{-14}
	algorithm in [11]	2.26* 10^{-06}	3.01* 10^{-06}
500	Theorem 4	6.84* 10^{-15}	1.41* 10^{-14}
	algorithm in [11]	1.20* 10^{-06}	1.53* 10^{-06}
700	Theorem 4	8.25* 10^{-15}	2.23* 10^{-14}
	algorithm in [11]	1.59* 10^{-06}	2.07* 10^{-06}
1000	Theorem 4	1.01* 10^{-14}	1.21* 10^{-14}
	algorithm in [11]	9.23* 10^{-05}	1.18* 10^{-05}

Example 2. Let $s = 1$, $a = 1$ and $b = 2$, we test the algorithm based on Theorem 1 (Corollary 2) for the problem (2) with $X^T = X$ and $XX^T = X + 2I_n$ constraint. The test matrix A with singular values are set as follows,

$$\begin{aligned} &[U, \mathrm{temp}] = \mathrm{qr}(1 - 2 * \mathrm{rand}(n));\\ &[V, \mathrm{temp}] = \mathrm{qr}(1 - 2 * \mathrm{rand}(n));\\ &d = \mathrm{rand}(n, 1);\\ &A = U \mathrm{diag}(d) V^T; \end{aligned}$$

and the matrix B are constructed by the following rules,

$$\begin{aligned} &c(1 : n/2) = 2; c(n/2 + 1 : n) = -1\\ &[U, \mathrm{temp}] = \mathrm{qr}(1 - 2\mathrm{rand}(n));\\ &X_* = U \mathrm{diag}(c) U^T;\\ &B = AX_*; \end{aligned}$$

so the optimal residual will be zero.

We still let numbers $m = n$ be variant from 100 to 1000. For different n, the residual precision $\|AX - B\|_F$ can reach 10^{-13}. The symmetric error $\|X^T - X\|_F$ is zero always, and the error $\|XX^T - X - 2I_n\|_F$ can reach 10^{-14}. In Table 2, we list the CPU time, $\|AX - B\|_F$, and $\|XX^T - X - 2I_n\|_F$, respectively. Since $\|X^T - X\|_F$ are zero always, we omit it.

Example 3. In this experiment, we test the efficiency of our algorithms when the coefficient matrices have different condition numbers. We ask for the least squares problem (2) with symmetric idempotent constraint ($X^T = X$,and $XX^T = X$) by Theorem 1,

Table 2: Solve the problem (2) with $X^T = X$ and $XX^T = X + 2I_n$ constraints based on Theorem 1.

$n = m$	CPU(s)	$\|AX - B\|_F$	$\|XX^T - X - 2I_n\|_F$
100	0.14	2.57*10^{-14}	9.91*10^{-14}
200	0.45	5.25*10^{-14}	1.87*10^{-14}
300	0.95	7.91*10^{-13}	3.03*10^{-14}
500	1.24	1.62*10^{-13}	1.73*10^{-14}
700	3.20	1.89*10^{-13}	7.94*10^{-14}
1000	9.36	2.68*10^{-13}	1.29*10^{-13}

that is $s = 1$, $a = 1$, $b = 0$. The test matrix A with singular values are set as follows,

$$\begin{aligned}
&[U, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&[V, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&d = [1 + \text{rand}(9 * n/10, 1);\\
&\qquad 10^{-\alpha} * (\text{rand}(n/10, 1) + 0.1)];\\
&A = U\text{diag}(d)V^T;
\end{aligned}$$

and, the matrix B are constructed by the following rules,

$$\begin{aligned}
&c = \text{rand}(n, 1) > 0.5;\\
&[U, \text{temp}] = \text{qr}(1 - 2\text{rand}(n));\\
&X_* = U\text{diag}(c)U^T;\\
&B = AX_*;
\end{aligned}$$

where $\alpha > 0$ is a constant that determines the magnitudes of the condition number of A. In Table 3, we list the experiment results for given $n = 500$. CPU time almost remains unchangeable even the condition number of A become bigger. The symmetric error $\|X^T - X\|_F$ keeps zeros always. The other items listed are similar to the above example.

Table 3: Variant condition numbers for the problem (2) with $X^T = X$ and $XX^T = X$ constraints, $n = 500$

Cond(A)	CPU(s)	ϵ	$\|XX^T - X\|_F$
1.94*10^{2}	5.37	1.95*10^{-12}	8.47*10^{-14}
1.87*10^{3}	5.46	1.65*10^{-12}	8.53*10^{-14}
1.84*10^{4}	5.45	2.11*10^{-12}	8.96*10^{-14}
1.41*10^{5}	5.44	2.01*10^{-12}	7.77*10^{-14}
1.73*10^{6}	5.48	1.71*10^{-12}	7.87*10^{-14}
1.92*10^{7}	5.38	1.45*10^{-12}	7.65*10^{-14}
1.49*10^{7}	5.33	1.61*10^{-12}	7.34*10^{-14}
1.66*10^{8}	5.51	1.72*10^{-12}	6.68*10^{-14}
1.59*10^{9}	5.42	1.65*10^{-12}	6.65*10^{-14}
1.83*10^{10}	5.39	1.55*10^{-12}	8.97*10^{-14}
1.15*10^{11}	5.51	1.21*10^{-12}	8.78*10^{-14}

Example 4. Finally we test our algorithm for the least squares problem of $AX = B,\ XC = D$ with

$$PX = XP,\ X^T = X \text{ and } XX^T = I_n$$

constraints. The symmetric matrix P with two different eigenvalues is generated as follows:

$$\begin{aligned}
&[H, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&d = [repmat(1, [1, n/2]),\\
&\qquad repmat(4, [1, n/2])];\\
&P = H\text{diag}(d)H^T.
\end{aligned}$$

The coefficient matrix $A,\ C \in \mathbb{R}^{n\times n}$ are constructed by

$$\begin{aligned}
&[U_A, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&[V_A, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&d_A = \text{rand}(n, 1);\\
&A = U_A\text{diag}(d_A)V_A^T;\\
&[U_C, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&[V_C, \text{temp}] = \text{qr}(1 - 2 * \text{rand}(n));\\
&d_C = \text{rand}(n, 1);\\
&C = U_C\text{diag}(d_C)V_C^T;
\end{aligned}$$

and $B,\ D$ are constructed by

$$B = AH\text{diag}(X_1, X_2)H^T,$$

and

$$D = H\text{diag}(X_1, X_2)H^TC,$$

where $X_i,\ i = 1, 2$ are symmetric orthogonal matrices generated by following rule:

$$\begin{aligned}
&[U_{X_i}, \text{temp}] = \text{qr}(1 - 2\text{rand}(n/2));\\
&d_{X_i} = 1 - 2\text{rand}(n/2, 1);\\
&X_i = U_{X_i}\text{diag}(\text{sign}(d_{X_i}))U_{X_i}^T.
\end{aligned}$$

The numerical results in Table 4 show that our algorithm is effective for this constrained problem.

6 Conclusion

In this paper, we consider the balanced Procrustes problem with $X = sX^T$ and $XX^T = aX + bI_n$ for given numbers $a,\ b$. By one time eigenvalue decomposition or real Schur decomposition of the matrix product generated by the matrices A and B, we construct the constrained solutions simply. We also generalize these conclusions to the problem with corresponding P-commuting constraints with given symmetric matrix P and the extended equations $AX = B$, $XC = D$ with the same constraints.

Table 4: The $PX = XP,\ X^T = X$ and $XX^T = I_n$ constrained solutions to the least squares problem of $AX = B,\ XC = D$

n	CPU(s)	$\left\| \begin{pmatrix} B-AX \\ D-XC \end{pmatrix} \right\|_F$	$\|X^T - X\|_F$
100	0.15	$1.49*10^{-14}$	$2.23*10^{-14}$
200	0.79	$3.78*10^{-14}$	$1.43*10^{-14}$
400	5.36	$6.35*10^{-14}$	$1.65*10^{-14}$
600	15.45	$7.39*10^{-14}$	$1.78*10^{-14}$
800	37.11	$2.55*10^{-13}$	$1.54*10^{-14}$
n	CPU(s)	$\|XX^T - I_n\|_F$	$\|PX - XP\|_F$
100	0.15	$2.43*10^{-14}$	$2.42*10^{-13}$
200	0.79	$5.64*10^{-14}$	$3.54*10^{-13}$
400	5.36	$1.34*10^{-14}$	$3.66*10^{-13}$
600	15.45	$5.53*10^{-14}$	$5.68*10^{-13}$
800	37.11	$8.89*10^{-13}$	$3.64*10^{-12}$

Acknowledgements: The authors are grateful to referee and Professor Zhenyue Zhang (Department of Mathematics, Zhejiang University, Yuquan Campus, Hangzhou, 310027, P. R. China) for their enlightening suggestions. Moreover, the research was supported in part by the Natural Science Foundation of Zhejiang Province and National Natural Science Foundation of China (Grant Nos. Y6110639, 11201422)..

References:

[1] L. E. Andersson and T. Elfving, A constrained Procrustes problem, *SIAM J. Matrix Anal. Appl.*, 18(1), 1997, pp. 124-139.

[2] A. W. Bojanczyk and A. Lutoborski, The Procsrustes Problem for Orthogonal Stiefel Matrices, *SIAM J. Sci. Comp.*, 21(4), 1999, pp. 1291-1304.

[3] F. Crosilla and A. Beinat, Use of generalized Procrustes analysis for the photogrammetric block adjustment by independent models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(3), 2002, pp. 195-209.

[4] L. Eldén and H. Park, A Procrustes problem on the Stiefel manifold, *Numerische Mathematik*, 82(4), 1999, pp. 599-619.

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[6] J. C. Gower, Generalized Procustes analysis, *Psychometrika*, 40(1), 1975, pp. 33-51.

[7] J. C. Gower and G. B. Dijksterhuis, *Procrustes problems*, Oxford University Press, New York, 2004.

[8] N. J. Higham, The symmetric Procrustes problem, *BIT*, 28, 1988, pp. 133-143.

[9] U. Kintzel, Procrustes problems in finite dimensional indefinite scalar product spaces, *Linear Algebra Appl.*, 402, 2005, pp. 1-28.

[10] J. Kiskiras and G. D. Halikias, A note on the complex semi-definite matrix Procrustes problem, *Numer. Linear Algebra Appl.*, 14, 2007, pp. 485-502.

[11] C. J. Meng, X. Y. Hu and L. Zhang, The skew-symmetric orthogonal solutions of the matrix equation $AX = B$, *Linear Algebra Appl.*, 402, 2005, pp. 303-318

[12] C. J. Meng and X. Y. Hu, The inverse problem of symmetric orthogonal matrices and its optimal approximation, *Mathematica Numerica Sinica*, 28(3), 2006, pp. 269-280.

[13] Z. Y. Peng and X. Y. Hu, The reflexive and anti-reflexive solutions of the matrix equation $AX = B$, *Linear Algebra Appl.*, 375, 2003, pp. 147-155.

[14] P. H. Schönemann, A Generalized Solution of The Orthogonal Procrustes Problem, *Psychometrika*, 31(1), 1966, pp. 1-10.

[15] Y. Qiu and A. Wang, Solving balanced Prorutes problem with some constraints by eigenvalue decomposition, *Journal of Computational and Applied Mathematics*, 233, 2010, pp. 2916-2924.

[16] J. ten Berge and K. Nevels, A general solution to Mosier's oblique procrustes problem, *Psychometrika*, 42(4), 1977, pp. 593-600.

[17] Q. Wang, S. Yu and C. Li, Extreme ranks of a linear quaternion matrix expression subject to triple quaternion matrix equations with applications, *Applied Mathematics and Computation,* 195, 2008, pp. 733-744.

[18] Q. Wang, H. Chang and C, Lin, P-(skew) symmetric common solutions to a pair of quaternion matrix equations, *Applied Mathematics and Computation*, 195, 2008, pp. 721-732.

[19] Q. Wang, J. W. van der Woude and H. Chang, A system of real quaternion matrix equations with applications, *Linear Algebra and its Applications*, 431, 2009, pp. 2291-2303.